Biomarkers of stroke recovery: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable (SRRR).


Lara A. Boyd. Department of Physical Therapy & the Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver V6T1Z3, Canada.
Kathryn S. Hayward. Department of Physical Therapy, University of British Columbia, Vancouver V6T1Z3, Canada; Stroke Division, The Florey Institute of Neuroscience and Mental Health, Heidelberg 3084, Australia; NHMRC Centre of Research Excellence in Stroke Rehabilitation and Brain Recovery, Australia.
Nick S. Ward. Sobell Department of Motor Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK.
Cathy M. Stinear. Department of Medicine and Centre for Brain Research, University of Auckland, Auckland New Zealand.
Charlotte Rosso. Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, F-75013, France; AP-HP, Stroke Unit, Pitié-Salpêtrière Hospital, 75013, France.
Rebecca J. Fisher. Division of Rehabilitation & Ageing, University of Nottingham, UK
Alexandre R. Carter. Department of Neurology, Washington University in Saint Louis, St Louis 63130, United States.
Alex P. Leff. Department of Brain Repair and Rehabilitation, Institute of Neurology & Institute of Cognitive Neuroscience, University College London, Queens Square WC1N 3BG, UK.
David A. Copland. School of Health & Rehabilitation Sciences, University of Queensland, Brisbane 4072 Australia; and University of Queensland Centre for Clinical Research, Brisbane 4029 Australia.
Leanne M. Carey. School of Allied Health, College of Science, Health and Engineering, La Trobe, University, Bundoora 3086, Australia; and Stroke Division, The Florey Institute of Neuroscience and Mental Health, Heidelberg 3084, Australia.
Leonardo G Cohen. Human Cortical Physiology and Neurorehabilitation Section, NINDS, NIH
Jane M. Maguire. Faculty of Health, University of Technology, Sydney 2007, Australia; Hunter Medical Research Institute, Priority Research Centre for Stroke and Brain Injury, Newcastle 2305, Australia.
D. Michele Basso. The Ohio State University, School of Health and Rehabilitation Sciences, Columbus 43210, United States.
Keywords (3-7 total): Biomarker, motor, somatosensory, cognition, language, stratification, function, clinical trial, neuroimaging, human

List of tables and figures:
Table 1: Scope of functional domains considered given existing literature.
Table 2: Scope for expert consensus biomarker recommendations.
Table 3: Expert consensus biomarker recommendations.
Abstract

The most difficult clinical questions in stroke rehabilitation are “What is this patient’s potential for recovery?” and “What is the best rehabilitation strategy for this person, given her/his neurological profile?” Without answers to these questions, clinicians struggle to make decisions regarding whether therapy should focus on recovery or compensation, and researchers design studies that inadvertently mix individuals who have a high likelihood of responding with those who do not. Developing and implementing biomarkers that distinguish patient subgroups will help address these issues and unravel the factors important to the recovery process. The goal of the present paper is to provide a consensus statement regarding the current state of the evidence for stroke recovery biomarkers. Biomarkers of motor, somatosensory, cognitive and language functions across the recovery timeline post-stroke are considered. We provide evidence for biomarkers that are considered ready to be included in clinical trials, and others that are promising but not as well understood, which represent a developmental priority. We will provide an example that illustrates the utility of biomarkers in recovery and rehabilitation research, which demonstrates how the inclusion of biomarkers will enhance future clinical trials. In this way, we propose a way forward for when and where we can include biomarkers to advance the efficacy of rehabilitation and recovery trials.
Introduction

Stroke is a heterogeneous condition making prediction of outcome and treatment response difficult. Despite this, clinical trials are often designed with a 'one size fits all' point of view, which makes them vulnerable to patient heterogeneity and failure. Biomarkers may assist with patient selection for trials. A stroke recovery biomarker (SRB) can be defined as an indicator of disease state that can be used as a measure of underlying molecular/cellular processes that may be difficult to measure directly in humans, and could be used to understand outcome, or predict recovery or treatment response (Bernhardt et al., 2016). In practical terms biomarkers might: 1) improve our ability to predict long-term outcomes across multiple domains, and 2) identify who and when to target with interventions for promoting stroke recovery (Cramer, Koroshetz, & Finklestein, 2007).

Interventions fall into two broad mechanistic categories: 1) behavioural interventions that take advantage of experience and learning dependent plasticity (e.g. motor, speech and language therapy), and 2) treatments that enhance the potential for experience and learning dependent plasticity to maximise the effects of behavioural interventions (e.g. pharmacotherapy or non-invasive brain stimulation). To identify who and when to intervene, we need biomarkers that reflect the underlying biological mechanisms that are being targeted.

Our goal was to provide a consensus statement regarding the evidence for stroke recovery biomarkers that are helpful in outcome prediction and therefore identifying subgroups for stratification, in four functional domains (motor, somatosensation, cognition, language, Table 1) according to recovery phase post stroke (hyperacute: <24hrs; acute: 1 to 7 days; early subacute: 1 week to 3 months; late subacute: 3 months to 6 months; chronic: > 6 months (Bernhardt et al., 2016 in review). For each domain, we provide recommendations for biomarkers that are: 1) ready to guide stratification of subgroups of patients for clinical trials and/or to predict outcome, or 2) a developmental priority (Table 2). Finally, we provide an example of how inclusion of a clinical trial-ready biomarker might have benefitted a recent phase III trial. We recognize that technology is rapidly advancing and emphasize brain biomarkers of structure and function rather than method. As there is generally limited evidence at this time for blood or genetic biomarkers, we do not discuss
these but recognise they are a developmental priority (Farr & Wegener, 2010; Kim et al., 2016; Lindgren & Maguire, 2016; Whiteley et al., 2012).

A challenge across the reviewed literature is to determine where biomarker data explains recovery beyond that denoted by a clinical measure alone using the proportional recovery rule shown using both motor (Krakauer & Marshall, 2015) and language (Lazar et al., 2010) outcomes. Further, to fully understand the predictive capacity of biomarkers we need to move beyond cross-sectional studies, which constitute the bulk of the biomarker literature at present, and conduct longitudinal research that provides the data that can predict outcome or treatment response. This is a priority for the field.

**Motor**

Neuroimaging biomarkers include quantitative characterisation of the stroke lesion itself, as well as the structure and function of non-lesioned brain areas. There is little consensus regarding the usefulness of characterizing the ischemic penumbra at the **hyper-acute** stage. Recent data suggest that the site of ischemic penumbra, rather than volume, could predict outcome or treatment response (i.e. to thrombolysis) and affect motor recovery (Rosso & Samson, 2014). **Acute** infarct volume is also related to motor outcomes (Helenius & Henninger, 2015; Kasner, 2006).

The extent of existing white matter disease (i.e. leukoaraiosis) has been associated with **acute** lesion size, degree of lesion expansion and stroke severity measured by initial National Institute of Health Stroke Severity (NIHSS) score (Henninger et al., 2012). Periventricular white matter hyperintensities are associated with functional independence measure (FIM) motor scores at the **sub-acute** stage (Liou et al., 2010; Senda et al., 2016). Though the extent of non-lesion white matter disease relates to initial motor deficits, it is not known how white matter damage affects subsequent recovery. This area of research is a developmental priority.

Measures of the corticospinal tract (CST) in the **acute** stage may predict motor outcome. Early measurement of CST fiber number predicts motor outcome (Fugl-Meyer score) at 12 months, especially for patients with more severe initial impairment (Bigourdan et al., 2016). Corticospinal tract lesion load (Feng et al., 2015) also has a positive predictive value of 100% for poor motor
outcome (Fugl-Meyer score at 3 months). Importantly, this model improved prediction beyond what could be determined by Fugl-Meyer upper limb score, age and stroke volume alone. Other regions distant from the lesion influence motor recovery, including ipsilesional CST distal to the stroke lesion (Feng et al., 2015). Fractional anisotropy (FA) of the CST in the peduncle and posterior limb of the internal capsule at the **acute** stage, is higher in individuals who make good and moderate recovery than those who make poor recovery (Wen et al., 2016). At the **chronic** stage, greater damage to CST (assessed with either transcranial magnetic stimulation (TMS) or diffusion tensor imaging (DTI)) is associated with greater motor impairment (Kou et al., 2013; Schulz et al., 2012; Shelton & Reding, 2001). The structural integrity of non-lesioned structures, such as the corpus callosum, is related to motor impairment at the **chronic** stage (Mang et al., 2015). Other data suggest that distant regions including FA in the contralesional CST (Borich, Mang, & Boyd, 2012), precentral gyrual thickness (Borich, Neva, & Boyd, 2015), and FA in the superior longitudinal fasciculi (Buch et al., 2012; Censor, Buch, Nader, & Cohen, 2016) play a role in motor outcome and ability to learn at the **chronic** stage. Multivariate machine learning methods have recently been applied to neuroimaging data with the aim of providing individual predictions based on an approach that integrates features extracted from brain ‘voxels’ from multiple brain areas, rather than focus on one area e.g. CST (Mah). Such approaches have recently suggested the importance of taking account of damage in brain regions other than CST in accounting for variation in motor outcome (Park et al, 2016; Rondina et al 2016).

There is consensus that the presence of an upper limb motor evoked potential (MEP) in response to TMS at the **hyperacute** and **acute** stages strongly predicts good motor outcome (D’Olhaberriague et al., 1997; Delvaux et al., 2003; Pennisi et al., 1999; Rapisarda, Basting, de Noordhout, Pennisi, & Delwaide, 1996; Vang, Dunbabin, & Kilpatrick, 1999) (Stinear, Barber, Petoe, Anwar, & Byblow, 2012; van Kuijk, Pasman, Hendricks, Zwarts, & Geurts, 2009) and that shorter MEP latencies and central motor conduction times are associated with better outcome (Heald, Bates, Cartlidge, French, & Miller, 1993). The presence of a MEP also identifies which patients will follow the proportional recovery rule (Byblow, Stinear, Barber, Petoe, & Ackerley, 2015). In the leg, the presence of a MEP (MEP+) indicates that an individual is more likely to be independently mobile 12 months post-stroke (Chang, Do, & Chun, 2015; Piron, Piccione, Tonin, & Dam, 2005; Steube,
Wietholter, & Correll, 2001), yet this measure does not relate to walking recovery. Prediction of recovery is more challenging for patients without a MEP (Hayward et al., 2017), and combining TMS with MRI biomarkers may be useful here (Stinear et al., 2012). TMS at the chronic stage helps explain the relationship between corticomotor function and motor performance in cross-sectional studies, and those who are MEP+ are more likely to benefit from physical interventions (Koski, Mernar, & Dobkin, 2004; Lai et al., 2015; Stinear et al., 2007).

Resting state FC (rsFC) findings in the early and late subacute converge on the conclusion that interhemispheric connectivity is behaviorally important. There is a positive association between acute/early subacute rsFC measured in the ipsilesional primary motor cortex (M1) with the contralesional thalamus, supplementary motor area (SMA), and medial frontal gyrus and motor outcomes at 6-months (FM scores) (C. H. Park et al., 2011). In late subacute patients, the amount of CST damage combined with interhemispheric M1 rsFC predicts greater therapy-induced gains for non-lacunar stroke (Burke Quinlan et al., 2015). In those with lacunar stroke rsFC between ipsilesional M1 to ipsilesional premotor (PMd) predicts therapy-related behavioural gain (Burke Quinlan et al., 2015). Baseline rsFC can predict the response to robotic therapy (Fan et al., 2015), yet more work is necessary before rsFC can be declared a useful predictive biomarker for response to therapy. As rsFC can be performed in severe patients, can interrogate all brain networks simultaneously and is relatively simple to acquire, it represents a developmental priority.

Quantitative indices extracted from functional MRI (fMRI) in the early and late subacute stage such as the laterality index from M1, and the study of its change over time, show a less lateralized pattern of activation after stroke (Feydy et al., 2002; Schaechter et al., 2002a). One unifying conclusion across studies is that the best motor outcomes are associated with the greatest shift towards the normal state of brain function (Calautti & Baron, 2003). The laterality index has been used as judgment criterion of efficacy in the chronic stage in trials testing mirror therapy (Michielsen et al., 2011), constraint-induced therapy (Schaechter et al., 2002b) and robotic intervention protocols (Milot et al., 2014). As there are fewer long-term studies of laterality index this is an area of developmental priority.
Magnetoencephalography (MEG) or electroencephalography (EEG), non-invasive measures of cortical neuronal oscillations, are sensitive to alterations in both GABAergic and glutamatergic signalling that are important for plasticity and recovery after stroke (Paggiaro et al., 2016; Rabiller, He, Nishijima, Wong, & Liu, 2015; Ward, 2015). Changes in cortical excitation and inhibition represent novel therapeutic targets, but cannot be directly measured in humans. Stroke patients with poorer outcomes have persistent increased low-frequency oscillations at the acute, and early/late subacute stages (Laaksonen et al., 2013), suggesting predominant inhibitory mechanisms in peri-lesional cortex. Acutely, lower beta-rebound in response to tactile finger stimulation (which indicates increased early post-stroke sensorimotor excitability) (Laaksonen et al., 2012) and increased somatosensory map size (Roiha et al., 2011) predict good recovery after stroke. Lastly, in a single stroke patient, zolpidem reversed increased peri-lesional theta (4-10Hz) and beta oscillations leading to clinical improvement (Hall et al., 2010). While MEG/EEG cannot be recommended to guide stratification of subgroup in trials at present, this is a developmental priority.

Combining neuroimaging and neurophysiology biomarkers may be useful for predicting motor outcomes and therapy response (Kim and Weinstein 2016). Upper limb outcomes at 3-months can be predicted at the early sub-acute stage by combining clinical, MRI and TMS biomarkers in a stepdown approach, the PREP algorithm (Stinear et al., 2012). Stoykov & Stinear (2010) treated chronic stroke patients using active-passive bilateral arm training and discovered that fractional anisotropy asymmetry between the two CST tracts accounted for 40% variability in clinical improvement. Factoring in whether patients were MEP+/- improved the predictive model. In a similar approach, multiple baseline clinical/radiological measures were assessed to predict gains made during a robotic therapy; a lower level of impairment and task related motor cortex activation measured with fMRI were associated with lesser gains with therapy (Takahashi, Der-Yeghiaian, Le, Motiwala, & Cramer, 2008). The integrity of corticospinal fibres from M1 has also been found to correlate with response to robot-assisted therapy (Riley et al., 2011). More recently, the same group argue that CST damage, interhemispheric connectivity between motor cortices and the presence or absence of cortical injury were the key factors in determining response to 3 weeks of a robotic upper limb training intervention (Burke-Quinlan et al. 2016)
In summary, neuroimaging and neurophysiology CST biomarkers can predict motor outcome and response to therapy, and can be considered for use in clinical trials. The evidence for rsFC and fMRI biomarkers are promising and are an area of developmental priority Table 3A.

**Sensory**

Currently there are few studies of structural or functional biomarkers conducted to understand outcome, predict recovery or predict treatment response in the somatosensory functioning domain in the **hyperacute or acute** phases post-stroke. Most work for structural biomarkers involving non-lesioned brain has focused on understanding outcome by mapping the structural integrity of residual pathways. Feasibility of visualisation of sensorimotor systems by tracking fibres has been demonstrated in **hyperacute, acute and early subacute phases** for motor and somatosensory symptoms (Yamada et al., 2003). There are changes in morphology of the somatosensory cortex of chronic stroke patients (Schaechter, Moore, Connell, Rosen, & Dijkhuizen, 2006), with co-localized structural (cortical thickness) and functional (brain activation, tactile stimulation) effects. Only one study noted associations between the structure of somatosensory regions and motor outcome in the **chronic** phase (Brodie et al. (2014).

Failure to activate the somatosensory cortex during median nerve stimulation in the **acute** stage predicts poor clinical recovery at 3 months (Manganotti et al., 2012). Using MEG, reduction in interhemispheric asymmetries of activity at **chronic** compared to **acute** phases was associated with a worse clinical state (Tecchio et al., 2006). Studies using MEG in **early** and **late subacute** phases show that changes in the source strengths of the primary motor and somatosensory cortices correlate with the extent of recovery of sensorimotor functions as determined by neurological exams (Huang et al., 2004). In the **subacute** phase differences in brain activity measured with task-related fMRI correlated with touch impairment in patients with thalamus /internal capsule lesions compared to those with lesions of primary (SI) or secondary (SII) somatosensory cortex (Carey et al., 2011). Similarly, responsiveness of SI at 1-15 days post-stroke is associated with improvement of two-point discrimination 3 months post-stroke (Wikstrom et al., 2000). Resting-state FC studies of touch
impairment and recovery demonstrated changes in connectivity from contralesional SII and thalamus associated with improved touch sensation in the chronic stage (Bannister et al., 2015).

Associations are observed between somatosensory function and a ratio of fibers in the sensory component of the superior thalamic radiation at the chronic phase (A. Borstad, Schmalbrock, Choi, & Nichols-Larsen, 2012) and the frontoparietal tracts in the acute (Meyer et al., 2016) and chronic (A. L. Borstad, Choi, Schmalbrock, & Nichols-Larsen, 2016) phases. In addition, somatosensation function in the chronic phase correlates with activity in the ipsilesional and contralesional primary sensorimotor cortex (Jang & Lee, 2013) and a more distributed pattern of activity involving parietal cortex (A. Borstad et al., 2012). Improvement in touch discrimination at 6 months was associated with increased rsFC between seeds in the contralesional hemisphere and distributed regions, including cerebellum (Bannister et al., 2015). Using MEG, involvement of ipsilesional primary hand representation areas positively contributed to clinical recovery (Tecchio et al., 2007).

Changes have also been reported in association with training of touch discrimination (n = 15; Carey et al., 2016), passive proprioception (n = 15; Dechaumont-Palacin, et al. 2008) and sensorimotor function (n=2) (n = 2; Borstad et al, 2013); with focus on tracking outcomes and mechanisms, rather than prediction. For example, touch discrimination training of patients with somatosensory loss at 6 months post-stroke was associated with different patterns of change in activation with thalamic/capsular compared to SI/SII cortical lesions (Carey et al., 2016). This area of research is a developmental priority.

There is insufficient evidence to recommend the use of any specific biomarkers of somatosensory system function in clinical trials. The recovery of somatosensation is often overlooked despite well-documented observations that impaired sensation is an impediment to optimal motor recovery (Blennerhassett, Matyas, & Carey, 2007; A. L. Borstad & Nichols-Larsen, 2014; Kong, Chua, & Lee, 2011). Functional biomarkers, including task-related activation and rsFC are a developmental priority (Table 3B).

Cognition
There is no biomarker work in the **hyperacute or acute phase** that helps understand outcome, or predict recovery or treatment response for executive functioning after stroke. This is a significant limitation in the field, and may reflect the slow trajectory of recovery from cognitive impairments.

Impairments in abstract reasoning and executive functioning in the **early subacute phase** independently predict long-term cognitive impairment (6 to 10-months post stroke) (Nys et al., 2005). Mapping executive functions to specific brain regions is problematic because these functions are distributed widely across the brain and their relationships are complex. Indeed, studies that readily identify structure-function relationships for phonology and semantic processing, often fail to find an equivalent for executive function (Butler, Lambon Ralph, & Woollams, 2014). The most consistent relationships are found in white matter. In the lesioned area, frontal and basal ganglia region microbleeds are associated with executive dysfunction outcome in the **chronic phase** (Werring et al., 2004), while in the non-lesioned areas mean diffusivity of normal appearing white matter (whole brain) correlates with executive function outcome in individuals with ischaemic leukoaraiosis (plus a previous lacunar stroke) (O'Sullivan et al., 2004; Senda et al., 2016).

Though functional imaging methods may offer the best hope of generating robust biomarkers for executive function there is little published work. In the **late subacute phase** outcome on executive function tests correlates with alpha band functional connectivity between the left fronto-opercular cortex and the rest of the brain (Dubovik et al., 2013). Yet, it is possible that the task-dependent changes observed with functional imaging data has less to do with ‘new’ domain-specific areas being generated, and more to do with cognitive control networks improving residual performance (F. Geranmayeh, Brownsett, & Wise, 2014). In individuals with **chronic** post-stroke aphasia a positive correlation between task-dependent activity in midline frontal cortex and language recovery was interpreted as reflecting domain-general cognitive control systems (Brownsett et al., 2014). Perhaps we should be developing therapeutic techniques to train executive function; data from controls suggests that this approach is feasible (Glass, Maddox, & Love, 2013).

The default mode network (DMN) has emerged as a key for cognitive functioning (Fox et al., 2005; Raichle, 2015; Shulman et al., 1997). Studies in the **subacute** and early **chronic** phases report
altered rsFC in the DMN correlated with cognitive performance after stroke (Dacosta-Aguayo et al., 2015; Ding et al., 2014; J. Y. Park et al., 2014; Tuladhar et al., 2013) Re-emergence of the anticorrelation between the DMN and task-positive networks, such as the dorsal attention network (DAN) (Corbetta & Shulman, 2002), is associated with behavioural recovery of cognitive functions. Resting state studies of the DAN provide the most robust example of disruption of interhemispheric connectivity associated with domain-specific cognitive deficits (He et al., 2007) (Carter et al., 2010) (Baldassarre et al., 2014; Siegel, Ramsey, & Snyder, 2016) and recovery (Ramsey & Siegel, 2016). While correlational analyses cannot establish causality, the finding that a change in rsFC correlates with behaviour lends support to the idea that measures of network connectivity may be a useful and responsive biomarker across multiple behavioural domains.

Based on this evidence we are not yet at the stage where we have biomarkers that are ready for clinical trials of cognitive function, especially executive functioning. Thus, the identification of biomarkers that explain cognitive function is a significant developmental priority area. Resting state FC is a promising candidate biomarker (Table 3C), and study of its utility as a biomarker of recovery should be emphasized.

Language

There are a number of studies identifying a relationship between lesion site and aphasia (Plowman, Hentz, & Ellis Jr., 2012). In the hyperacute period, perfusion-weighted MRI showed that word comprehension deficits are strongly correlated with blood flow in Wernicke’s area (Hillis et al., 2001). A related study demonstrated that lexical processing was more strongly related to the volume of hypo-perfused tissue than the volume of lesion (Hillis, Barker, Beauchamp, Gordon, & Wityk, 2000). Imaging illustrates that recovery of word comprehension from the hyperacute to acute phase (3 days) is associated with reperfusion of Wernicke’s area (Hillis & Heidler, 2002). Recovery of naming in the hyperacute period is predicted by reperfusion of left posterior middle temporal/fusiform gyrus, Broca’s area, and/or Wernicke’s area (Croquelois, Wintermark, Reichhart, Meuli, & Bogousslavsky, 2003; Hillis et al., 2006; Reineck, Agarwal, & Hillis, 2005). There are no established predictors of long-term (>3-days) recovery from outcomes observed in the hyperacute period (<24 hours), thus
this is a developmental priority.

Impaired repetition in the **acute** phase was associated with structural damage to the arcuate fasciculus and Broca’s area and tissue dysfunction (hypoperfusion and frank damage) in the inferior portion of the left supramarginal gyrus and temporal-parietal junction (Fridriksson, 2010). Impaired repetition was associated with posterior temporal-parietal lesions and damage to the dorsal superior longitudinal and arcuate fasciculus, while comprehension deficits were associated with ventral extreme capsule fibre damage (Kummerer et al., 2013). A multivariate machine learning technique using a mask of task-induced fMRI activity in bilateral frontal and temporal regions was used in combination with behavioural language performance and age. This approach correctly predicted language recovery in 86% of individuals with stroke who had aphasia at 2 weeks (Saur et al., 2006).

In the early **subacute phase**, there are relationships between lesion location and aphasia symptoms. Analyses based on lesion location and symptoms correctly classified 67% to 94% of patients. Forkel et al. (2014) demonstrated predictions of recovery at 6 months were improved by adding volume of the left long segment of the arcuate fasciculus to a regression model including age, sex, and lesion size. Including volume of the right long segment of the arcuate fasciculus further improved recovery prediction. Recent work by Geranmayeh et al. (2016) showed that propositional language production is predicted by interactions between brain networks (default mode network, fronto-temporo-parietal, and cingulo-opercular networks) rather than activity within individual networks, highlighting the distributed nature of language operations.

Voxel-based analyses in the **chronic** phase have identified structural damage associated with particular aphasic symptoms, distinguishing between semantic and phonological processes and recognition versus production (Bates et al., 2003; Dronkers et al., 2004; Mirman et al., 2015; Buttler et al. 2014; Leff et al., 2009). Arcuate fasciculus lesion load negatively influences speech production (Marchina et al., 2011) and classifies severe and non-severe outcomes with 90% accuracy for naming and 96% accuracy for speech fluency (Wang, Marchina, Norton, Wan, & Schlaug, 2013). The PLORAS (Predicting Language Outcome and Recovery After Stroke) system (Price, Seghier, & Leff, 2010) uses a Gaussian process model regression with a large database of stroke patients (from 1 month post, therefore covering early and late subacute, and chronic phases) with structural MRI,
demographic, and language performance to provide predictions of aphasia recovery at the individual level. Using this approach and covariate factors of time of stroke, volume, and 35 different brain regions, predictions of language outcome, and within subject changes in speech production, have been identified (Hope, Seghier, Leff, & Price, 2013).

Posterior middle temporal damage can negatively affect aphasia therapy outcome in the chronic phase (Fridriksson, 2010). Meinzer et al. (2010) observed a negative relationship between the proximity of the lesion to the hippocampus and response to a naming treatment. Bonilha et al. (2016) showed that measures of neural network connectivity combined with initial severity accounted for 78% of variance in response to anomia treatment. Several small studies identified a relationship between therapy success and integrity of the left arcuate fasciculus (van Hees et al., 2014), the right arcuate fasciculus (Schlaug, Marchina, & Norton, 2009) and white matter in proximity to the hippocampus (Meinzer et al., 2010). Further, several fMRI studies have investigated treatment-induced aphasia recovery, predominantly in the chronic stage. Fridriksson (2010) identified a significant relationship between treatment-induced naming improvements and fMRI activity in a both a posterior cluster (including parietal lobe and precuneus) and an anterior cluster (including middle frontal gyrus and pars opercularis). Subsequent analyses (Fridriksson, Richardson, Fillmore, & Cai, 2012) showed that altered activity in perilesional areas was associated with increased naming accuracy, but measures of pre-treatment brain activity (as opposed to changes in activity) predicted improvement in semantic errors, suggesting additional factors contribute to treatment outcome.

In summary, in the acute and early subacute stages, the use of structural MRI and DTI provide insights into the neural basis of language deficits, but there are not sufficient large studies demonstrating that these methods improve prediction of recovery or treatment response. fMRI shows potential at the early subacute stage for significantly improving prediction of outcome (Saur et al., 2010), however, this approach needs validation. Structural MRI and DTI may forecast recovery at the late subacute and chronic stage, suggesting the possible use of these techniques to stratify patients for clinical trials, understand therapy mechanism and predict outcome. It should be noted that: (1) there is still considerable variability in outcome that is not accounted for by these methods, (2) each
method uses a unique and complex analysis technique, and (3) different aphasia treatments may engage unique networks (Table 3D).

**Conclusions**

How might biomarker data be incorporated into future research? The term “stroke” is inadequate, as it describes a very heterogeneous group that is unified by a vascular injury, but not by size, location, impact, or injury context. Biomarkers present a way forward to subgroup or stratify patients to reduce variance and increase power, allowing for smaller sample sizes (Cramer, 2010). Moreover, the final behavioral phenotype after stroke can arise from many different biological states, which could result in differential therapeutic responses. A patient exploiting all possible compensatory brain mechanisms might have little room to improve, while a similar patient who uses no compensatory mechanisms might achieve treatment benefits (Hardwick, Rajan, Bastian, Krakauer, & Celnik, 2017). Only with the inclusion of a biomarker can we attempt to disentangle recovery vs. compensation.

Clinical trials therefore need to base participant eligibility on more than stroke status or behavioural measures. Patient selection must include biomarkers; ideally these will be linked with preclinical methods as well as the biological mechanism of the therapy or treatment under investigation. For example, recently a threshold was defined whereby patients in the early and late subacute stage with >63% injury to the corticospinal tract did not achieve clinically significant gains associated with a robotic therapy (Burke Quinlan et al., 2015). This result highlights the ascendant role that neuroimaging measures need to play in clinical-decision making for post-stroke rehabilitation (Menon, Campbell, Levi, & Goyal, 2015).

A useful example comes from the recent phase III Everest trial (Levy et al., 2016), which relied on behavioural assessments to determine participant eligibility, and ultimately found that patients randomized to epidural stimulation did not reach the primary efficacy endpoint more often than patients in the control group. However, a post hoc analysis of patients randomized to epidural stimulation found that the primary efficacy endpoint was reached more often (67%) by those with preserved motor evoked responses upon cortical stimulation compared to those lacking a response.
Thus, had confirmation of physiological integrity of the biological target been an eligibility criterion, the effect size would have been substantially higher and the trial results quite different. We believe that this example is highly useful in illustrating the utility of biomarkers in recovery and rehabilitation research and expect that the inclusion of biomarkers will enhance future clinical trials.
Acknowledgements: We would like to acknowledge Prof Julie Bernhardt for convening this first Stroke Recovery and Rehabilitation Roundtable (SRRR), Ms Katie Lafferty for moderating discussions at SRRR and Dr Karen Borschmann for project management. We thank Prof Leonid Churilov for his assistance with scoping survey design and analysis. The Florey Institute of Neuroscience and Mental Health was the coordinating centre, for which we acknowledge support from the Victorian Government, in particular from the Operational Infrastructure Support Grant. Financial support was provided by NHMRC CRE Stroke Rehabilitation & Brain Recovery (Australia), Heart and Stroke Foundation of Canada, and Heart and Stroke Foundation Canadian Partnership for Stroke Recovery. An unrestricted educational grant was provided by Ipsen Pharma.
References


